Cooperative Wideband Spectrum Sensing over Fading Channels

Hongjian Sun, *Member, IEEE,* Arumugam Nallanathan, *Senior Member, IEEE,* Shuguang Cui, *Senior Member, IEEE,* and Cheng-Xiang Wang*, *Senior Member, IEEE*

***Abstract*—In cognitive radio (CR) systems, it is crucial for secondary users to reliably detect spectral opportunities across a wide frequency range. This paper studies a novel multi-rate sub-Nyquist spectrum sensing (MS^3) system capable of performing wideband spectrum sensing in a cooperative CR network over fading channels. The aliasing effects of sub-Nyquist sampling are modelled. To mitigate such effects, different sub-Nyquist sampling rates are applied such that the numbers of samples at different CRs are consecutive prime numbers. Moreover, the performance of MS^3 over fading channels (Rayleigh fading and log-normal fading) is analysed in the form of bounds on the probabilities of detection and false alarm. The key finding is that the wideband spectrum can be sensed using sub-Nyquist sampling rates in MS^3 over fading channels, without the need of spectral recovery. In addition, the aliasing effects can be mitigated by the use of different sub-Nyquist sampling rates in a multi-rate sub-Nyquist sampling system.**

***Index Terms*—Cognitive radio, wideband spectrum sensing, sub-Nyquist sampling, Rayleigh fading, log-normal fading.**

I. Introduction

A crucial requirement of cognitive radios (CRs) is that they must be able to rapidly find and make good use of spectral opportunities without causing harmful interference to the primary user (PU) [1], [2]. The ability of finding spectral opportunities is called spectrum sensing, which is considered as one of the most critical components in a CR system. When the frequency range is sufficiently narrow such that the channel frequency response can be considered flat, narrowband spectrum sensing algorithms can be applied , e.g., matched-filtering, cyclostationary feature detection, and energy detection [3]. Horgan and Murphy [4] analysed the performance of energy detection over Nakagami-m fading channels. Sofotasios *et al.* [5] investigated the case of applying energy detection over generalised $\kappa-\mu$ and $\kappa-\mu$ extreme fading channels. The analysis was then extended to the case of cooperative spectrum sensing where each CR can transmit its decision or measurement to a fusion center (FC) where final decision is made. A simple fusion scheme used in the FC is hard decision fusion which fuses decisions from multiple CRs. It is upper bounded by the soft decision fusion scheme (i.e., square-law combining scheme [6]) where innite precision energy vectors are transmitted to the FC. Herath *et al.* [7] and Atapattu *et al.* [8], [9] further analysed the performance of energy detection on the signal processed by maximal ratio combining (MRC), equal gain combining (EGC), and selection combining (SC) schemes over Nakagami-m and Rician fading channels together with shadowing effects.

Ideally we hope that, if a PU reappears, CRs have several other possible vacant subbands to access, facilitating a seamless hand-off from one spectral channel to another. Unfortunately, aforesaid narrowband spectrum sensing algorithms ignore the diversified individual spectral opportunities across the wideband spectrum. Driven by the desire of exploiting wider bandwidth in CR networks, revolutionary wideband spectrum sensing techniques become increasingly important and deserve exploratory research. In previous work, Quan *et al.* [10], [11] proposed a multiband joint detection (MJD) approach that can sense the primary signal over a wide frequency range. It has been shown that MJD has superior performance for wideband spectrum sensing. In [12], Tian and Giannakis studied a wavelet detection approach, which could adapt parameters to a dynamic wideband spectrum. Furthermore, they elegantly applied the compressed sensing (CS) theory to implement wideband spectrum sensing by using sub-Nyquist sampling techniques in [13]. Later on, the CS-based approach has attracted much attention [14]–[19] owing to its advantage of using much fewer samples to perform wideband spectrum sensing. In our previous work [21], to save system energy, adaptive CS-based spectrum sensing approach was proposed that could find the best spectral recovery with high confidence. Unfortunately, using CS-based approaches, the spectral recovery requires high computational complexity, leading to a high spectrum sensing overhead that may be a serious issue in CRs with restricted computational resources.

The major contributions of this paper can be summarized as follows:

H. Sun is with the School of Engineering and Computing Sciences, the University of Durham, Durham, UK. (e-mail: hongjian@ieee.org)

A. Nallanathan is with the Department of Informatics, King's College London, London, WC2R 2LS, UK. (e-mail: nallanathan@ieee.org)

S. Cui is with the Department of Electrical and Computer Engineering, Texas A&M University (e-mail: cui@ece.tamu.edu). S. Cui is also a Distinguished Adjunct Professor at King Abdulaziz University in Saudi Arabia and a Visiting Professor at ShanghaiTech University, China.

C.-X. Wang (corresponding author) is with School of Information Science and Engineering, Shandong University, Jinan 250100, China, and the Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, U.K. (e-mail: chengxiang.wang@hw.ac.uk).

The research leading to these results has received funding from the European Commission's Horizon 2020 Framework Programme (H2020/2014-2020) under grant agreement No. 646470, SmarterEMC2 Project. The authors gratefully acknowledge the support of this work from EU FP7 QUICK project (No. PIRSES-GA-2013-612652), the 863 project in 5G wireless networking (No. 2014AA01A701), Ministry of Science and Technology of China, the Opening Project of the Key Laboratory of Cognitive Radio and Information Processing (Guilin University of Electronic Technology), Ministry of Education (No. 2013KF01). The work of S. Cui was supported in part by DoD with grant HDTRA1-13-1-0029, by NSF with grants CNS-1343155, ECCS-1305979, and CNS-1265227, and by grant NSFC-61328102.

- We introduce a multi-rate sub-Nyquist spectrum sensing (MS^3) approach for cooperative wideband spectrum sensing in a CR network. Since the spectral occupancy is low, sub-Nyquist sampling is adopted in each sampling channel to wrap the sparse spectrum occupancy map. The sensing overhead is therefore significantly reduced. The effects caused by sub-Nyquist sampling are analyzed, and the test statistic is represented by a reduced data set obtained from multi-channel sub-Nyquist sampling.
- We propose to use different sampling rates in different sampling channels (equivalently different CRs) for improving the spectrum sensing performance. Specifically, in the same observation time, the numbers of samples in different sampling channels are chosen as different consecutive prime numbers.
- We mathematically analyze the performance of MS^3 over fading channels, and then derive some closed-form bounds for the average probabilities of false alarm and detection.

The key advantage of MS^3 is that the wideband spectrum usages can be detected directly from a few sub-Nyquist samples without full spectral recovery that is usually required by other sub-Nyquist techniques. Compared with the existing spectrum sensing methods, MS^3 can achieve better wideband spectrum sensing performance with lower implementation complexity.

The rest of the paper is organized as follows. Section II introduces traditional spectrum sensing using energy detector. In Section III, we propose the wideband spectrum sensing approach, i.e., MS^3. The performance analysis of MS^3 fusing faded signals is given in Section IV. Section V presents simulation results, with conclusions given in Section VI.

II. Traditional Spectrum Sensing Using Energy Detection

Let us assume that all CRs keep quiet during the spectrum sensing interval as enforced by protocols, e.g., via the medium access control (MAC) layer [10]. Therefore, the observed spectral energy arises only from PUs and the background noise. Additionally, we assume that on each frequency bin at most one PU sends data, e.g., when the orthogonal frequency-division multiple access (OFDMA) transmission scheme is used by PUs. The total bandwidth of the signal sensed at each CR is W (Hertz). Over an observation time T, if the sampling rate f ($f \geq 2W$) is adopted, a sequence of Nyquist samples will be obtained with the length of fT. The observation time T is chosen such that fT is a non-prime natural number, e.g., $T = \frac{14}{f}$; thus, fT can be written as $fT = JN$, where both J and N are natural numbers. Furthermore, this length-JN sequence is divided into J equal-length segments and each segment has the length of N. If we use $x_{\mathrm{c},i}(t)$ ($t \in [0, T]$) to represent the continuous-time signal received at CR i, after Nyquist sampling, the sampled signal can be denoted by $x_i[n] = x_{\mathrm{c},i}(n/f),\ n = 0, 1, \cdots, JN-1$. At CR i, the sampled signal of segment j ($j \in [1, J]$) can be written as

$$x_{i,j}[n] = \begin{cases} x_{\mathrm{c},i}\left(\frac{(j-1)N+n}{f}\right), & n = 0, 1, \cdots, N-1 \\ 0, & \text{Otherwise.} \end{cases} \quad (1)$$

The discrete Fourier transform (DFT) spectrum of the sampled signal in segment j, i.e., $\overrightarrow{X_{i,j}}$, is given by

$$X_{i,j}[k] = \sum_{n=0}^{N-1} x_{i,j}[n] e^{-\jmath 2\pi kn/N},\ k = 0, 1, \cdots, N-1 \quad (2)$$

where $\jmath = \sqrt{-1}$. We model spectrum sensing on frequency bin k as a binary hypothesis test, i.e., $\mathcal{H}_{0,k}$ (absence of PU) and $\mathcal{H}_{1,k}$ (presence of PU) [11]:

$$X_{i,j}[k] = \begin{cases} Z_{i,j}[k], & \mathcal{H}_{0,k} \text{ or } k \in \Omega_i^{'} \\ H_{i,j}[k]S_{i,j}[k] + Z_{i,j}[k], & \mathcal{H}_{1,k} \text{ or } k \in \Omega_i \end{cases} \quad (3)$$

where $Z_{i,j}[k]$ is complex additive white Gaussian noise (AWGN) with zero mean and variance $\delta_{i,k}^2$, i.e., $Z_{i,j}[k] \sim \mathcal{CN}(0, \delta_{i,k}^2)$, $H_{i,j}[k]$ denotes the discrete frequency response between the PU and CR i, $S_{i,j}[k]$ is assumed to be a deterministic signal sent by the PU on frequency bin k, Ω_i denotes the spectral support such that $\Omega_i = \{k|\text{PU presents at } X_{i,j}[k]\}$, and $\Omega_i^{'} = \{k|\text{PU does not present at } X_{i,j}[k]\}$. Here, we consider that all CRs are sensing the same spectrum caused by the same set of PUs such that the sets $\Omega = \Omega_1 = \cdots = \Omega_i$. In addition, for simplicity, we assume that the noise variance $\delta_{i,k}^2$ is normalized to be 1. The observation time T is chosen to be smaller than the channel coherence time such that the magnitude of $H_{i,j}[k]$ remains constant within T at each CR, i.e., constant $|H_{i,j}[k]|$ over different segment j.

Since an energy detector does not require any prior information about the transmitted primary signal, with lower complexity than other spectrum sensing schemes [20], we adopt the energy detection approach in this paper. The received signal energy at each frequency bin can be calculated as

$$E_i[k] = \sum_{j=1}^{J} |X_{i,j}[k]|^2, \quad k = 0, 1, \cdots, N-1. \quad (4)$$

The decision rule at frequency bin k is then given by

$$E_i[k] \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\gtrless}} \lambda_k, \quad k = 0, 1, \cdots, N-1 \quad (5)$$

where λ_k is the detection threshold. The signal energy distribution on frequency bin k can be modeled as [20]

$$E_i[k] \sim \begin{cases} \chi_{2J}^2, & \mathcal{H}_{0,k} \text{ or } k \in \Omega_i^{'} \\ \chi_{2J}^2(2\gamma_i[k]), & \mathcal{H}_{1,k} \text{ or } k \in \Omega_i \end{cases} \quad (6)$$

where $\gamma_i[k] \triangleq \frac{|H_i[k]S_i[k]|^2}{\delta_{i,k}^2}$ denotes the SNR on the frequency bin k at CR i, χ_{2J}^2 denotes the central chi-square distribution, and $\chi_{2J}^2(2\gamma_i[k])$ denotes the non-central chi-square distribution. Both of these distributions have $2J$ degrees of freedom and $2\gamma_i[k]$ denotes the non-centrality parameter. The probabilities of false alarm and detection are given by [20]

$$P_{\mathrm{f},i,k}^{Nyq} = \Pr(E_i[k] > \lambda_k | \mathcal{H}_{0,k}) = \frac{\Gamma(J, \frac{\lambda_k}{2})}{\Gamma(J)} \quad (7)$$

$$P_{\mathrm{d},i,k}^{Nyq} = \Pr(E_i[k] > \lambda_k | \mathcal{H}_{1,k}) = Q_J\left(\sqrt{2\gamma_i[k]}, \sqrt{\lambda_k}\right) \quad (8)$$

where $\Gamma(a)$ denotes the gamma function, $\Gamma(a, x)$ denotes

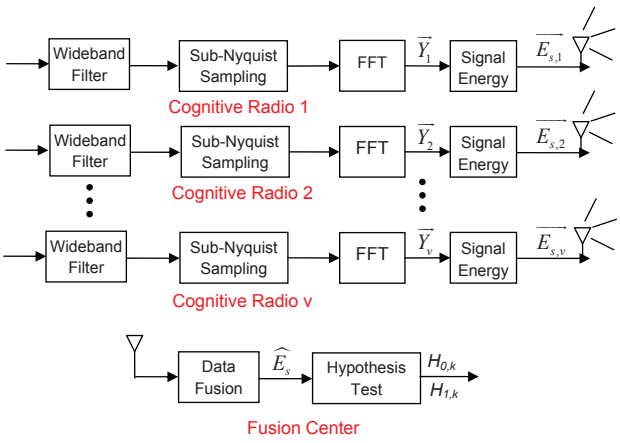

Fig. 1. Block diagram of multi-rate sub-Nyquist spectrum sensing system.

the upper incomplete gamma function, and $Q_u(a,x)$ is the generalized Marcum Q-function defined by $Q_u(a,x) = \frac{1}{a^{u-1}}\int_x^\infty t^u e^{-\frac{a^2+t^2}{2}} I_{u-1}(at)dt$ in which $I_v(a)$ is the v-th order modified Bessel function of the first kind.

III. Multi-rate sub-Nyquist spectrum sensing

It is difficult to realize wideband spectrum sensing, since it requires a high-speed Analog-to-Digital-Converter (ADC) for Nyquist rate sampling. We now present an MS^3 scheme using multiple low-rate sub-Nyquist samplers to implement wideband spectrum sensing in a CR network.

A. System Description

Consider that there are v synchronized CRs collaborating for wideband spectrum sensing, and the FC could be one of the CRs and have good channel gain with the other CRs (e.g., geographically nearby). Due to low primary spectral occupancy [13], the received signals at CRs are naturally sparse in the frequency domain. Here, we assume that the Nyquist DFT spectrum $\overrightarrow{X_{i,j}}$ defined in (2) is s-sparse ($s \ll N$), which implies that only the largest s out of N components need to be counted. The spectral sparsity level, i.e., s, can be obtained from sparsity estimation [16] or other methods. As shown in Fig. 1, MS^3 consists of several CRs, each of which has one wideband filter, one low-rate sampler, and a fast Fourier transform (FFT) device, where the wideband filters are set to have bandwidth of W. The operation of MS^3 can be described as follows:

1) As the coordinator of MS^3, the FC allocates different sub-Nyquist sampling rates to different CRs according to Theorem 1.
2) CRs perform sub-Nyquist samplings during the observation time T.
3) The sub-Nyquist DFT spectrum is calculated by using sub-Nyquist samples and an FFT device.
4) The signal energy vectors are formed by using the sub-Nyquist DFT spectrum.
5) The CRs transmit these signal energy vectors, i.e., $\vec{E}_{\mathrm{s},1}, \cdots, \vec{E}_{\mathrm{s},v}$ as calculated in (15), to the FC by using a dedicated common control channel.
6) The received data from all CRs is fused in the FC to form a test statistic.
7) The FC chooses the detection threshold and performs binary hypothesis tests.
8) The FC shares the detection results with all CRs via the dedicated control channel.

B. Sub-Nyquist Sampling and Data Combining

At CR i, we use sub-Nyquist rate f_i ($f_i < 2W \le f$) to sample the continuous-time signal $x_{\mathrm{c},i}(t)$. The sampled signal can be denoted by $y_i[n] = x_{\mathrm{c},i}(n/f_i),\ n = 0,1,\cdots,JM_i-1$ where $JM_i = f_iT$ and M_i is chosen to be a natural number. The sampled signal is then divided into J equal-length segments. The segment j ($j \in [1,J]$) can be written as

$$y_{i,j}[n] = \begin{cases} x_{\mathrm{c},i}\left(\frac{(j-1)M_i+n}{f_i}\right), & n = 0,1,\cdots,M_i-1 \\ 0, & \text{Otherwise.} \end{cases} \tag{9}$$

The DFT spectrum of the sampled signal of segment j ($j \in [1,J]$) is then given by

$$Y_{i,j}[m] = \sum_{n=0}^{M_i-1} y_{i,j}[n]e^{-\jmath 2\pi mn/M_i},\ m = 0,1,\cdots,M_i-1 \tag{10}$$

With the aid of Poisson summation formula [22], the DFT spectrum of sub-Nyquist samples can be represented by the DFT spectrum of Nyquist samples as proved in Appendix A:

$$Y_{i,j}[m] = \frac{M_i}{N}\sum_{l=-\infty}^{\infty} X_{i,j}[m+lM_i],\quad m = 0,1,\cdots,M_i-1 \tag{11}$$

where l is an unknown integer within $[0, N/M_i-1]$ such that $m+lM_i \in \Omega_i$.

One issue caused by sub-Nyquist sampling is signal overlapping in $Y_{i,j}[m]$. However, if we choose parameter N such that $N \gg s$ and let the sub-Nyquist sampling rate satisfy $M_i \sim \mathcal{O}(\sqrt{N})$, the probability of signal overlapping is very small (as proved in Appendix B). As such, we only focus on two cases: no signal on frequency bin m and one signal on frequency bin m. In the first case, only noise exists. In the second case, only a single l is active in (11), and the other terms in the summation of (11) can be modeled as noise based on (3). Thus, the following equation holds from (3) and (11):

$$Y_{i,j}[m] = \begin{cases} \frac{M_i}{N}\sum_\nu Z_{i,j}[m+\nu M_i], & \text{First ca} \\ \frac{M_i}{N}X_{i,j}[m+lM_i] + \frac{M_i}{N}\sum_{\nu\neq l} Z_{i,j}[m+\nu M_i], & \text{Second} \end{cases} \tag{12}$$

Furthermore, using (3) and (12), we can model the DFT spectrum distribution of sub-Nyquist samples by

$$\sqrt{\frac{N}{M_i}}Y_{i,j}\left[|k|_{\mathrm{mod}\,(M_i)}\right] \sim \begin{cases} \mathcal{CN}\left(0, \delta^2_{\mathrm{s},i,k}\right), & k \\ \mathcal{CN}\left(\sqrt{\frac{M_i}{N}}H_{i,j}[k]S_{i,j}[k], \delta^2_{\mathrm{s},i,k}\right), & k \end{cases} \tag{13}$$

where $\delta^2_{\mathrm{s},i,k}$ is the noise variance of sub-Nyquist DFT spectrum, which can be represented by the noise variance of

Nyquist DFT spectrum using (11):

$$\delta_{\mathrm{s},i,k}^2 = \underbrace{\left\lceil \frac{N}{M_i} \right\rceil}_{No.\ of\ sums} \left(\frac{M_i}{N} \underbrace{\sqrt{\frac{N}{M_i}}}_{Scaling\ of\ Y_{i,j}} \right)^2 \delta_{i,k}^2 \approx \delta_{i,k}^2 \quad (14)$$

where $\lceil \frac{N}{M_i} \rceil$ (the smallest integer not less than $\frac{N}{M_i}$) denotes the number of summations in (11).

The signal energy of sub-Nyquist DFT spectrum at each CR node is then calculated as

$$E_{\mathrm{s},i}[m] = \sum_{j=1}^{J} \left| Y_{i,j}[m] \right|^2, \quad m = 0, 1, \cdots, M_i - 1. \quad (15)$$

whose distribution can be modeled with (13) and (15) as

$$\frac{N}{M_i} E_{\mathrm{s},i} \left[|k|_{\bmod (M_i)} \right] \sim \begin{cases} \chi_{2J}^2, & k \in \Omega_i^{'} \\ \chi_{2J}^2 \left(2\frac{M_i}{N} \gamma_i[k] \right), & k \in \Omega_i, \end{cases} \quad (16)$$

where the length of energy vector $\vec{E}_{\mathrm{s},i}$ is M_i. We note that, due to the sub-Nyquist sampling, the noise is folded from the whole bandwidth onto all signals of interest as shown in (12). As a result, comparing (16) with (6), we find that the received SNR in the sub-Nyquist sampling channel i is degraded from γ_i to $\frac{M_i}{N}\gamma_i$. This SNR degradation depends on the ratio between the number of samples at the sub-Nyquist rate and that of at the Nyquist rate (i.e., $\frac{M_i}{N}$).

In MS^3, the signal energy vectors $\vec{E}_{\mathrm{s},1}, \cdots, \vec{E}_{\mathrm{s},v}$ calculated in (15) are then transmitted from CRs to the FC, which leads to a test statistic:

$$\widehat{E_{\mathrm{s}}}[k] = \sum_{i=1}^{v} \frac{N}{M_i} E_{\mathrm{s},i}[|k|_{\bmod (M_i)}], \quad k = 0, 1, \cdots, N-1 \quad (17)$$

where $\frac{N}{M_i}$ is a scaling factor to reconcile the different noise variance of CRs caused by different sub-Nyquist sampling rates. To test whether the PU is present or not, we adopt the following decision rule:

$$\widehat{E_{\mathrm{s}}}[k] \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\gtrless}} \lambda_k, \quad k = 0, 1, \cdots, N-1. \quad (18)$$

To analyse the performance of the decision rule in (18), we aim to model $\widehat{E_{\mathrm{s}}}$ through the use of (16) and (17). However, this is very challenging since different energy vectors in (17) may contain different mirror images of the original PU frequencies. To assist the analysis, we divide the set $\Omega_i^{'}$ of (16) into two disjoint subsets $\Omega_{\mathrm{A},i}^{'}$ and $\Omega_{\mathrm{U},i}^{'}$ such that $\Omega_{\mathrm{A},i}^{'} \cup \Omega_{\mathrm{U},i}^{'} = \Omega_i^{'}$ and $\Omega_{\mathrm{A},i}^{'} \cap \Omega_{\mathrm{U},i}^{'} = \emptyset$. Here, let $\Omega_{\mathrm{A},i}^{'}$ denote a set of aliased frequencies (i.e., false frequencies appear as mirror images of the original PU frequencies around the sub-Nyquist sampling frequency), and $\Omega_{\mathrm{U},i}^{'}$ represent a set of unaffected/unoccupied frequencies. Accordingly, the set $\Omega_{\mathrm{A},i}^{'}$ can be written as

$$\Omega_{\mathrm{A},i}^{'} \triangleq \left\{ k \middle| k = |m|_{\bmod (M_i)} + lM_i, \ \ k \in [0, N-1], m \in \Omega_i \right\} \quad (19)$$

where $|m|_{\bmod (M_i)}$ is used for describing the aliasing effect due to sub-Nyquist sampling and lM_i is used for accounting for the index extension from $0, \cdots, M_i - 1$ to $0, \cdots, N-1$ in (17). Consequently, the set $\Omega_{\mathrm{U},i}^{'}$ can be defined as

$$\Omega_{\mathrm{U},i}^{'} \triangleq \left\{ k \middle| k \notin \Omega_{\mathrm{A},i}^{'}, k \in \Omega_i^{'} \right\}. \quad (20)$$

Note that the aliased frequencies set $\Omega_{\mathrm{A},i}^{'}$ can lead to the same distribution as that of Ω_i due to the aliasing effect. Thus, we can model $\frac{N}{M_i}\vec{E}_{\mathrm{s},i}$ using (16) as

$$\frac{N}{M_i} E_{\mathrm{s},i}[k] \sim \begin{cases} \chi_{2J}^2, & k \in \Omega_{\mathrm{U},i}^{'} \\ \chi_{2J}^2 \left(2\frac{M_i}{N} \gamma_i[m] \right), & k \in \Omega_{\mathrm{A},i}^{'}, m \in \Omega_i \\ \chi_{2J}^2 \left(2\frac{M_i}{N} \gamma_i[k] \right), & k \in \Omega_i. \end{cases} \quad (21)$$

where $\gamma_i[m]$ denotes the SNR on the frequency bin m which is related to the observed frequency bin k by $k = |m|_{\bmod (M_i)} + lM_i$ as shown in (19). Such a representation is used for explaining the aliasing effect: observing a frequency bin k, we may find spectral component from another frequency bin m which folds back to k due to sub-Nyquist sampling.

Revisiting (17), we find that summing up energy vectors from different CRs will result in two new sets: $\Omega_{\mathrm{A}}^{'} \triangleq \cup_{i=1}^{v} \Omega_{\mathrm{A},i}^{'}$ and $\Omega_{\mathrm{U}}^{'} \triangleq \cap_{i=1}^{v} \Omega_{\mathrm{U},i}^{'}$, where $\Omega_{\mathrm{A}}^{'}$ denotes the set of aliased frequencies and $\Omega_{\mathrm{U}}^{'}$ is the set of unaffected/unoccupied frequencies of $\widehat{E_{\mathrm{s}}}$ due to the sum operation. The definition of the set $\Omega_{\mathrm{U}}^{'}$ is based on the fact that the frequency bin k of $\widehat{E_{\mathrm{s}}}$ can be classified as unaffected/unoccupied only if $k \in \Omega_{\mathrm{U},i}^{'}, \forall i \in [1, v]$. Furthermore, since $\widehat{E_{\mathrm{s}}}[k], k \in \Omega_{\mathrm{U}}^{'}$ is only affected by the noise, we can model it by the central chi-square distribution with $2Jv$ degrees of freedom. On the other hand, the definition of $\Omega_{\mathrm{A}}^{'}$ is due to the fact that the frequency bin k of $\widehat{E_{\mathrm{s}}}$ can be classified as aliased if $k \in \Omega_{\mathrm{A},i}^{'}$ in one or more than one channels (i.e., CRs). For convenience, let Υ_k denote the set of CRs in which each CR can generate aliased frequencies on the frequency bin k, i.e., $\Upsilon_k \triangleq \{i | k \in \Omega_{\mathrm{A},i}^{'}\}$. Additionally, applying the property of the sum of non-central chi-square variables to (17), we can model $\widehat{E_{\mathrm{s}}}[k], k \in \Omega_{\mathrm{A}}^{'}$ as a non-central chi-square distribution with $2Jv$ degrees of freedom and non-centrality parameter of $2 \sum_{i \in \Upsilon_k} \frac{M_i}{N} \gamma_i[m]$ by using (16). Based on above discussions, the distribution of $\widehat{E_{\mathrm{s}}}$ can be modelled as

$$\widehat{E_{\mathrm{s}}}[k] \sim \begin{cases} \chi_{2Jv}^2, & k \in \Omega_{\mathrm{U}}^{'} \\ \chi_{2Jv}^2 \left(\frac{2}{N} \sum\limits_{i \in \Upsilon_k} M_i \gamma_i[m] \right), & k \in \Omega_{\mathrm{A}}^{'}, m \in \Omega \\ \chi_{2Jv}^2 \left(\frac{2}{N} \sum\limits_{i=1}^{v} M_i \gamma_i[k] \right), & k \in \Omega \end{cases} \quad (22)$$

where both $k \in \Omega_{\mathrm{U}}^{'}$ and $k \in \Omega_{\mathrm{A}}^{'}$ mean that the frequency bin k is not used by PUs, thus corresponding to the hypothesis $\mathcal{H}_{0,k}$. The former case denotes that there are no aliased frequencies on the frequency bin k, while the latter case implies that there are $\mathrm{card}(\Upsilon_k)$ aliased frequencies on the frequency bin k. Here, $\mathrm{card}(\Upsilon_k)$ means the cardinality of the set Υ_k, i.e., the number of elements of the set Υ_k, which is equivalent to the number of CRs that have aliased frequencies on the frequency bin k. Thus, when analysing the probability of false alarm, we know that the former case leads to a lower probability of false alarm since $\widehat{E_{\mathrm{s}}}[k], k \in \Omega_{\mathrm{U}}^{'}$ contains noise only. In contrast, the latter

case $\widehat{E_{\rm s}}[k], k \in \Omega_{\rm A}^{'}$ leads to a higher probability of false alarm since it contains not only noise energy but also the energy of aliased frequencies due to sub-Nyquist sampling. Hence, the probability of false alarm on the frequency bin k can be bounded by considering two cases $k \in \Omega_{\rm U}^{'}$ and $k \in \Omega_{\rm A}^{'}$ in (22) and applying similar approach used in (6), (7), and (8):

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq P_{{\rm f},k} \leq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i\in\Upsilon_k} M_i\gamma_i[m]}, \sqrt{\lambda_k}\right). \quad (23)$$

C. Multi-rate Sub-Nyquist Spectrum Sensing

To improve the probability of false alarm, we shall reduce the number of elements in the set Υ_k in (23). This is mainly because the sum term $\sum\limits_{i\in\Upsilon_k} M_i\gamma_i[m]$ will decrease as the number of elements in the set Υ_k decreases, thus leading to a smaller upper bound on the probability of false alarm. Obviously, the card(Υ_k) is affected by sub-Nyquist sampling rates, since the card(Υ_k) describes how many CRs have aliased frequencies on the frequency bin k, which are originally caused by sub-Nyquist rate sampling. Using the same sub-Nyquist sampling rates in MS^3 is not recommended because it could lead to card(Υ_k)$= v$, resulting in a high probability of false alarm. Hence, we focus on applying different sub-Nyquist sampling rates in MS^3. For choosing sub-Nyquist sampling rates that can reduce the card(Υ_k), we start from the simplest case (for the purpose of analysis) where we only have one original PU frequency, i.e., only one frequency bin $k_1 \in \Omega$ is occupied by the PU and the spectral sparsity level $s = 1$.

Lemma 1: If the numbers of samples at multiple CRs, i.e., $M_1, M_2, ..., M_v$, are different primes, and meet the requirement of

$$M_iM_j > N, \quad \forall\, i \neq j \in [1, v], \quad (24)$$

two or more CRs cannot have mirrored frequencies on the same frequency bin.

The proof of Lemma 1 is given in Appendix C.

Second, considering the case where we have multiple original PU frequencies, we reach the following conclusion.

Lemma 2: If the numbers of samples at multiple CRs satisfy the conditions in Lemma 1, the number of elements in the set Υ_k will be bounded by the sparsity level s.

Proof: Based on the result of Lemma 1, only one CR can map the original frequency bin $k_j \in \Omega_i$ to the aliased frequency in $\Omega_{\rm A}^{'}$. Furthermore, recall $\Upsilon_k \triangleq \{i | k \in \Omega_{{\rm A},i}^{'}\}$ and the definition of $\Omega_{{\rm A},i}^{'}$ in (19), we know card(Υ_k) $\leq$ card(Ω_i) $= s$. □

Based on these two Lemmas, we quantify the detection performance of MS^3 as below.

Theorem 1: In MS^3, different sub-Nyquist sampling rates are adopted at different CRs such that the numbers of samples at CRs, i.e., $M_1, M_2, \cdots, M_v$, are different consecutive primes, and meet the requirement of $M_iM_j > N,\ \forall\ i \neq j \in [1, v]$. Using the decision rule of (18), we can obtain the probabilities of false alarm and detection with the following bounds:

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq P_{{\rm f},k} \leq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i\in\Upsilon_k, {\rm card}(\Upsilon_k)\leq s} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right) \quad (25)$$

$$P_{{\rm d},k} \geq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i=1}^{v} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right). \quad (26)$$

Proof: Using (23) and the bound card(Υ_k) $\leq s$, (25) follows. Furthermore, when the energy of one spectral component in Ω maps to another spectral component in Ω, the probability of detection will increase. Thus, the inequality of (26) holds. □

Remark 1: It is worth noting that the exact expression of the probabilities of false alarm and detection cannot be obtained in MS^3 systems. This is mainly caused by the aliasing effects of sub-Nyquist sampling, which unpredictably fold original frequencies back to different but unknown frequency bins. In detail, the set $\Omega_{{\rm A},i}^{'}$ in (21) is unknown and different from that of the other channel because of the sub-Nyquist sampling and unknown Ω_i (actually the purpose of spectrum sensing is to find it). Therefore, the set $\Omega_{\rm A}^{'}$ in (22) is typically unknown, making us impossible to know the exact expression of the probability of false alarm. This also occurs for analysing the probability of detection.

Remark 2: It is noteworthy to emphasize that the proposed sub-Nyquist system cannot achieve the same performance as the counterpart multi-CR Nyquist system due to its sub-Nyquist operations; however, it can obtain comparable spectrum sensing performance to a single-CR Nyquist system (at the same total sampling rate) with better flexibility and scalability. In particular, comparing (8) with (26), we find that the performance of MS^3 is comparable to the single-CR system when $\sum_{i=1}^{v} M_i \simeq N$ where " $\simeq$ " means approximately equal, given that:

$$P_{{\rm d},k} \geq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i=1}^{v} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right) \simeq Q_{Jv}\left(\sqrt{2\gamma_i[k]}, \sqrt{\lambda_k}\right)$$

$$Q_{Jv}\left(\sqrt{2\gamma_i[k]}, \sqrt{\lambda_k}\right) > Q_J\left(\sqrt{2\gamma_i[k]}, \sqrt{\lambda_k}\right) = P_{{\rm d},i,k}^{Nyquist} \quad (27)$$

and the probabilities of false alarm are similar when the sparsity level s is sufficiently small. Additionally, we note that MS^3 is more flexible and scalable than the single-CR Nyquist-sampling based detection system, since both the sub-Nyquist sampling rates at CRs and the number of CRs can be flexibly chosen to meet the condition $\sum_{i=1}^{v} M_i \simeq N$. Furthermore, we emphasize that MS^3 is applicable to the scenario of insufficient measurements, i.e., $\sum_{i=1}^{v} M_i < N$.

Remark 3: It can be seen from Theorem 1 that the sampling rates in MS^3 can be set much lower than the Nyquist rate since $M_i \sim \mathcal{O}(\sqrt{N})$. On the other hand, a higher average sampling rate can provide us with tighter bounds, since the probability of signal overlapping in the aliased spectrum can be reduced with a larger M_i in each sampling channel as discussed in Section III-B. Thus, there is a trade-off between the bound accuracy and the sampling rate saving. It should

be emphasized that there is no exact closed-form expression for the probabilities in Theorem 1, since the exact number of CRs that have aliased frequencies on the frequency bin k cannot be quantified. Moreover, we note that the upper and lower bounds in Theorem 1 can be easily computed since the Marcum-Q function can be efficiently computed using power series expansions [23]. Under the Neyman-Pearson criterion, we should design a test with the constraint of $P_{\mathrm{f},k} \leq \alpha$. In such a scenario, we must let the upper bound of (25) to be α and solve the detection threshold λ_k from the inverse of the Marcum-Q function. It has been shown in [24] that the detection threshold can be calculated with low computational complexity. In addition, to calculate the detection threshold, the noise power is required to be known at the FC.

IV. MS^3 over Fading Channels

In this section, we consider that the primary signals are propogated from PUs to CRs over fading channels, subject to either a Rayleigh or log-normal distribution. Recall the fusion rule in (17), we find that it is difficult to model the distribution of the sum of weighted independent random variables in (17) over fading channels. Hence, we use the sum of uniformly weighted random variables to approximate the sum of differently weighted random variables in Theorem 1:

$$\frac{2\sum_{i=1}^{v} M_i\gamma_i}{N} \simeq \frac{2\overline{M}}{N}\sum_{i=1}^{v}\gamma_i = \psi\gamma_{\mathrm{v}}, \qquad \frac{2\sum_{i\in\Upsilon_k} M_i\gamma_i}{N} \simeq \frac{2\overline{M}}{N}\sum_{i\in\Upsilon_k}\gamma_i = \psi\gamma_{\mathrm{s}} \tag{28}$$

where $\overline{M}$ is the average of M_i over multiple CRs, $\psi \triangleq \frac{2\overline{M}}{N}$, $\gamma_{\mathrm{v}} \triangleq \sum_{i=1}^{v}\gamma_i$, and $\gamma_{\mathrm{s}} \triangleq \sum_{i\in\Upsilon_k}\gamma_i$. We note that the above approximation accuracy mainly depends on $\frac{|\overline{M}-M_i|}{N}$, where a smaller $\frac{|\overline{M}-M_i|}{N}$ corresponds to a more accurate approximation. Since $M_1, M_2, ..., M_v$ are chosen to be v different consecutive prime numbers and the distance between primes could be very small compared to N, the parameter $\frac{|\overline{M}-M_i|}{N}$ approaches zero as N increases. Thus, the above approximation has little impact on the final result for large Ns.

A. Rayleigh Distribution

If the magnitudes of received primary signals at different CRs follow Rayleigh distributions, the SNRs will follow exponential distributions. Hence, γ_{v} and γ_{s} follow Gamma distributions:

$$f(\gamma_{\mathrm{v}}) = \frac{\gamma_{\mathrm{v}}^{v-1}}{\overline{\gamma}^{v}\Gamma(v)}e^{-\frac{\gamma_{\mathrm{v}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{v}} \geq 0, \qquad f(\gamma_{\mathrm{s}}) = \frac{\gamma_{\mathrm{s}}^{s-1}}{\overline{\gamma}^{s}\Gamma(s)}e^{-\frac{\gamma_{\mathrm{s}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{s}} \geq 0 \tag{29}$$

where $\overline{\gamma} = \mathbb{E}(\frac{|HS|^2}{\delta^2})$ denotes the average SNR over multiple CRs, and $f(\cdot)$ denotes a generic probability density function (PDF) of its argument.

The bounds for the average probabilities of false alarm and detection in MS^3 can be obtained by averaging (25) and (26) over all possible SNRs, respectively.

Theorem 2: If the magnitudes of received signals at different CRs follow Rayleigh distribution, the average probabilities of false alarm ($\overline{P_{\mathrm{f},k}}$) and detection ($\overline{P_{\mathrm{d},k}}$) in MS^3 will have the following bounds

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq \overline{P_{\mathrm{f},k}} \leq \Theta(s, Jv, \psi, \overline{\gamma}[k], \lambda_k) \tag{30}$$

$$\overline{P_{\mathrm{d},k}} \geq \Theta(v, Jv, \psi, \overline{\gamma}[k], \lambda_k) \tag{31}$$

where $\Theta(x, Jv, \psi, \overline{\gamma}, \lambda)$ is defined as

$$\Theta = \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=0}^{\infty}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n}\frac{\Gamma\left(n+Jv, \frac{\lambda}{2}\right)}{\Gamma\left(n+Jv\right)} \tag{32}$$

in which C_a^b denotes the binomial coefficient, i.e., $\mathrm{C}_a^b = \frac{b!}{a!(b-a)!}$.

The proof of Theorem 2 is given in Appendix D.

Remark 4: From Theorem 2, we have $0 \leq \Theta \leq 1$, since the term $\frac{\Gamma(a,b)}{\Gamma(a)} \in [0,1]$ and the remaining terms can be simplified to 1. In addition, it can be proved that Θ is a monotonically increasing function with respect to ψ, $\overline{\gamma}$, and x, respectively. Therefore, both probabilities will either increase or remain the same when the average sampling rate and the average SNR increase, more sampling channels will lead to a higher probability of detection, and the average probability of false alarm can be reduced with smaller s.

Remark 5: Since (32) contains infinite sums, its computational complexity is directly related to the number of computed terms, which are required to obtain a specific accuracy. As the number of computed terms (i.e., P) varies, the magnitude of the truncation error can be written as

$$\begin{aligned} T_\Theta(P) &= \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=P}^{\infty}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n}\frac{\Gamma\left(n+Jv, \frac{\lambda}{2}\right)}{\Gamma\left(n+Jv\right)} && (33)\\ &\leq \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=P}^{\infty}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} && (34)\\ &= 1-\left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=0}^{P-1}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} && (35)\end{aligned}$$

where the inequality of (34) holds for $\frac{\Gamma(n,\frac{\lambda}{2})}{\Gamma(n)} \leq 1$, and (35) is obtained by using the binomial expansion. It can be shown that (32) converges very quickly. For example, in order to achieve double-precision accuracy, only $P = 30 \sim 40$ calculated terms are required; therefore, the bounds are tractable. To solve for the detection threshold λ_k, we could use the lower bound on $P_{\mathrm{f},k}$ in (31). This is due to the fact that the lower bound can approximate $P_{\mathrm{f},k}$ very well as analyzed in Appendix D.

B. Log-normal Distribution

The strength of the transmitted primary signal may also be affected by shadowing from buildings, hills, and other objects. A common model is that the received power fluctuates with a log-normal distribution. In such a scenario, the PDF of the SNR at CR i, i.e., $f(\gamma_i)$, is given by

$$f(\gamma_i) = \frac{\xi}{\sqrt{2\pi}\sigma_i\gamma_i}\exp\left(-\frac{\left(10\log_{10}(\gamma_i)-\overline{\gamma}_i\right)^2}{2\sigma_i^2}\right), \quad \gamma_i > 0 \tag{36}$$

where $\xi = 10/\ln(10)$, and σ_i (dB) denotes the standard deviation of $10\log_{10}\gamma_i$ at CR i. Note that the PDF in (36) can

be closely approximated by a Wald distribution [20], [25]:

$$f(\gamma_i) = \sqrt{\frac{\eta_i}{2\pi}}\gamma_i^{-3/2}\exp\left(-\frac{\eta_i(\gamma_i-\theta_i)^2}{2\theta_i^2\gamma_i}\right), \quad \gamma_i > 0 \quad (37)$$

where $\theta_i = \mathbb{E}(\gamma_i)$ denotes the expectation of γ_i, and η_i is the shape parameter for CR i. Via the method of moments, the parameters η_i, θ_i and $\overline{\gamma}_i$, σ_i are related as follows:

$$\theta_i = \exp\left(\frac{\overline{\gamma}_i}{\xi} + \frac{\sigma_i^2}{2\xi^2}\right), \qquad \eta_i = \frac{\theta_i}{\exp(\frac{\sigma_i^2}{\xi^2}) - 1}. \quad (38)$$

In the proposed system, the condition $\frac{\eta_i}{\theta_i^2} = \frac{\mathbb{E}(\gamma_i)}{\mathrm{Var}(\gamma_i)} = b$ (constant) can be satisfied. Thus, γ_s and γ_v also follow the Wald distribution [26]. The PDFs of γ_s and γ_v are given by

$$f(\gamma_\mathrm{s}) = \sqrt{\frac{s\eta}{2\pi}}\gamma_\mathrm{s}^{-3/2}\exp\left(-\frac{\eta(\gamma_\mathrm{s}-s\theta)^2}{2s\theta^2\gamma_\mathrm{s}}\right), \quad \gamma_\mathrm{s} > 0 \quad (39)$$

$$f(\gamma_\mathrm{v}) = \sqrt{\frac{v\eta}{2\pi}}\gamma_\mathrm{v}^{-3/2}\exp\left(-\frac{\eta(\gamma_\mathrm{v}-v\theta)^2}{2v\theta^2\gamma_\mathrm{v}}\right), \quad \gamma_\mathrm{v} > 0 \quad (40)$$

where η and θ denote the averages of η_i and θ_i, respectively.

Theorem 3: If the magnitudes of received signals at different CRs follow log-normal distributions, the average probabilities of false alarm ($\widetilde{P_{\mathrm{f},k}}$) and detection ($\widetilde{P_{\mathrm{d},k}}$) in MS^3 will be bounded as

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \le \widetilde{P_{\mathrm{f},k}} \le \Lambda(s, Jv, \psi, \lambda_k, \theta[k], \eta[k]) \quad (41)$$

$$\widetilde{P_{\mathrm{d},k}} \ge \Lambda(v, Jv, \psi, \lambda_k, \theta[k], \eta[k]) \quad (42)$$

where $\Lambda(x, Jv, \psi, \lambda, \theta, \eta)$ is defined by

$$\Lambda = \sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=0}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\Gamma\left(n+Jv,\frac{\lambda}{2}\right)}{n!\Gamma\left(n+Jv\right)}\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}\mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(x\psi\theta^2+\eta)}{\theta^2}}\right) \quad (43)$$

in which $\mathrm{K}_{n-\frac{1}{2}}(a)$ denotes the modified Bessel function of the second kind with order $n-\frac{1}{2}$.

The proof of Theorem 3 is given in Appendix E.

Remark 6: Because (43) contains infinite sums, the truncation error $T_\Lambda(P)$ must be considered. Similar to (34), the truncation error can be written as

$$\begin{aligned} T_\Lambda(P) &\le \sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=P}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!}\mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(x\psi\theta^2+\eta)}{\theta^2}}\right) \\ &= 1-\sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=0}^{P-1}\frac{\left(\frac{\psi}{2}\right)^n\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!}\mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(x\psi\theta^2+\eta)}{\theta^2}}\right) \end{aligned} \quad (44)$$

V. Simulation Results and Analysis

In our simulations, we assume that the CRs are organized as shown in Fig. 1 and adopt the following configurations unless otherwise stated. We use the wideband analog signal model in [27] and thus the received signal $x_{\mathrm{c},i}(t)$ at CR i has the form:

$$x_{\mathrm{c},i}(t) = \sum_{l=1}^{N_b}|H_{i,l}|\sqrt{E_l}B_l\cdot\mathrm{sinc}\left(B_l(t-\Delta)\right)\cdot\cos\left(2\pi f_l(t-\Delta)\right)+z(t) \quad (45)$$

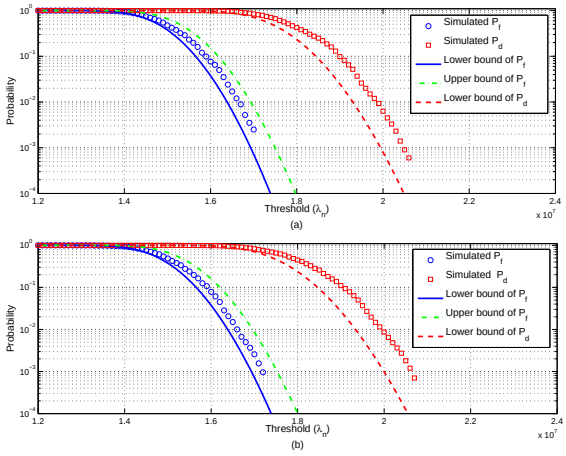

Fig. 2. Comparisons of simulation results and theoretical bounds on the probabilities of false alarm and detection in Theorem 1 and Theorem 2, respectively, when MS^3 combining (a) non-faded signals, (b) Rayleigh faded signals with the received SNR= 5 dB at CRs.

where N_b denotes the number of non-overlapping subbands, $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, Δ denotes a random time offset, $z(t)$ is AWGN, i.e., $z(t) \sim \mathcal{N}(0,1)$, E_l is the transmit power at PU, and $H_{i,l}$ denotes the discrete frequency response between the PU and CR i in subband l. We assume that there are $v = 22$ CRs in the CR network and the channels from PUs to CRs are independent and subject to fading. The received signal $x_{\mathrm{c},i}(t)$ consists of $N_b = 6$ non-overlapping subbands. The l-th subband is in the frequency range of $[f_l - \frac{B_l}{2}, f_l + \frac{B_l}{2}]$, where the bandwidth $B_l = 1 \sim 10$ MHz and f_l denotes the center frequency. The center frequency of the subband l is randomly located within $[\frac{B_l}{2}, W - \frac{B_l}{2}]$ (i.e., $f_l \in [\frac{B_l}{2}, W - \frac{B_l}{2}]$), where the overall signal bandwidth $W = 10$ GHz. In MS^3, the received signal is sampled by using different sub-Nyquist rates at different CRs for $T = 20$ μs. To be specific, the numbers of samples at multiple CRs are chosen following Theorem 1 and we choose the first prime $M_1 \approx a\sqrt{N}$ ($a \ge 1$) along with its $v-1$ neighboring and consecutive primes in the increasing direction. The spectral observations are obtained by applying an FFT to these sub-Nyquist samples in each channel. Then the signal energy is calculated in the spectral domain using (15), and the energy vectors are transmitted from the CRs to the FC using dedicated common control channels. In the FC, we form the test statistic with (17). We define the compression rate as the ratio between the number of samples at the sub-Nyquist rate and the number of samples at the Nyquist rate, i.e., $\frac{M}{N}$, where M denotes the average number of sub-Nyquist samples at CRs. Spectrum sensing results are obtained by using the decision rule (18) and varying the detection threshold λ_k.

In Fig. 2, we verify the theoretical results in (25)-(26) and (30)-(31) by comparing these bounds with the simulated results. It shows that the bounds on the probabilities of false alarm and detection can predict the simulated results. Fig. 2 also illustrates that the lower bound on the probability of detection can successfully predict the trend of simulated results. Fig. 3 shows the receiver operating characteristic (ROC) curves of

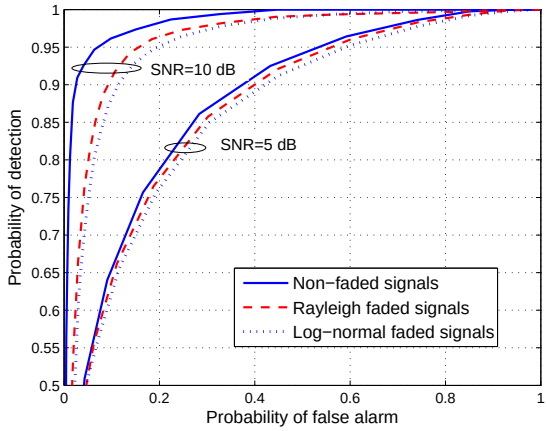

Fig. 3. Receiver operating characteristic curves of MS^3 for combining non-faded signals or faded signals when the compression rate $\frac{M}{N} = 0.0219$ and the number of segments $J = 5$. The wideband signal is observed by 22 CRs at different sampling rates (the average sampling rate is 448.68 MHz).

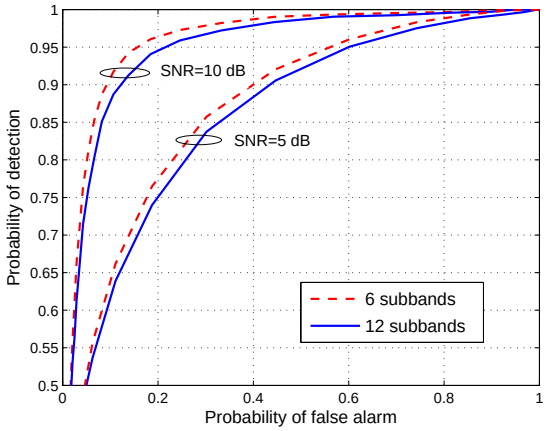

Fig. 4. The performance of MS^3 for combining Rayleigh faded signals with $v = 22$ and $\frac{M}{N} = 0.0219$, when the received SNR at CRs and the number of subbands change.

MS^3 when combining non-faded and faded signals. When the average SNR as received at CRs is 5 dB, the performance of MS^3 combining faded signals is roughly the same as that of combining non-faded signals, since the strength of the signal is mostly masked by the noise. In contrast, the detection performance of MS^3 combining non-faded signals outperforms that of combining faded signals when SNR=10 dB. In addition, we see that the performance of MS^3 combining log-normal shadowed signals is the poorest. Nonetheless, even for log-normal shadowed signals, MS^3 has a probability of nearly 90% for detecting the presence of PUs when the probability of false alarm is 10%, with the compression rate of $\frac{M}{N} = 0.0219$.

To investigate the influence of s and SNR, we use Fig. 4 to show the performance of MS^3 when the received signals are faded according to Rayleigh distribution with different values of s (proportional to the number of subbands). We see that, as the number of subbands decreases, the detection performance improves for the same SNR. The performance improvement of MS^3 stems from that, for a fixed number of sampling channels, decreasing s makes it easier to distinguish the occupied frequencies from the aliased frequencies as discussions in Section III-C.

In Fig. 5, we compare the performance of MS^3 against that of Nyquist systems when all systems have similar probabilities of false alarm. In the Nyquist system type I, each CR is given an orthogonal subband (wideband spectrum is divided into several equal-length subbands) to sense at Nyquist rate, while their decisions are sent back to the FC. In the Nyquist system type II, we assume that each CR must measure the whole wideband spectrum independently, thus requiring multiple standard ADCs in each node to cover all wideband spectrum. After signal sampling, all measurements are sent back to the FC, where equal gain combining approach is adopted to fuse data and then energy detection is used for spectrum sensing. Fig. 5 shows that the proposed system has superior performance to the Nyquist system type I. On the other hand, it can also be seen that the Nyquist system type II has performance gain over

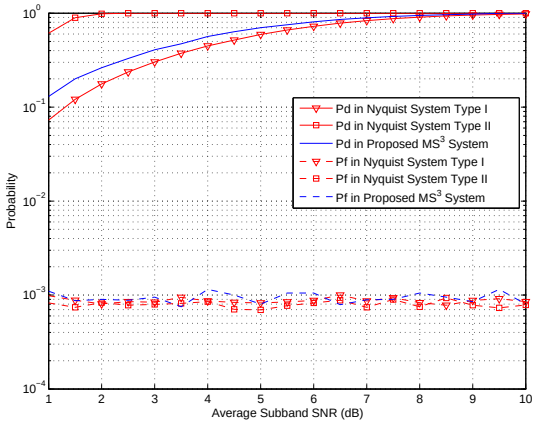

Fig. 5. The probability of detection (P_d) and the probability of false alarm (P_f) of the proposed MS^3 system and Nyquist systems over log-normal shadowing channels when the standard deviation $\sigma = 5$ dB and the average subband SNR varies. In the proposed MS^3 system, the wideband signal is sampled at different sampling rates by 22 ADCs with the average sampling rate of 448.68 MHz, and the compression rate is $\frac{M}{N} = 0.0219$.

the proposed system; however, at the expense of much higher implementation complexity as discussed below.

In Table I, we compare the implementation complexity of MS^3 with that of the Nyquist systems, when the received signals at different CRs are faded according to Rayleigh distribution. Here, we consider the following comparison metric: the number of same-sampling-rate ADCs for achieving $P_{\rm d} \geq 90\%$ and $P_{\rm f} \leq 10\%$, since practical CRs often have requirements on the probabilities of detection and false alarm to secure the performance of both CRs and PUs. We see that, when there exist 10 CRs, MS^3 requires each CR to be equipped with a single ADC with an average sampling rate of 957.54 MHz; thus, the whole CR network only requires 10 low-rate ADCs. In contrast, the Nyquist system type I requires 21 ADCs in total, because of 21×957.54 MHz$\approx$ 20 GHz for covering 10 GHz spectrum based on Nyquist sampling theorem. In the Nyquist system type II, 210 ADCs (with the average sampling

TABLE I
IMPLEMENTATION COMPLEXITY COMPARISON OF MS^3 AND THE NYQUIST SYSTEMS WHEN THE RECEIVED SIGNALS ARE FADED ACCORDING TO RAYLEIGH DISTRIBUTION WITH TEN DECIBEL RECEIVED SNR AT CRS.

Number of CRs for Wideband Spectrum Sensing (v)	10	20	30	40
Required Number of ADCs in the Proposed MS^3 System	10	20	30	40
Required Number of ADCs in the Nyquist System Type I	21	40	58	74
Required Number of ADCs in the Nyquist System Type II	21×10	40×20	58×30	74×40
Average Sampling Rate of ADCs in the Above Systems ($\overline{f}$ in MHz)	957.54	513.08	350.34	276.77
Sample Reduction Rate I ($\frac{\text{Total samples in MS}^3}{\text{Total samples in Nyquist type I}}$)	47.62%	50%	51.72%	54.05%
Sample Reduction Rate II ($\frac{\text{Total samples in MS}^3}{\text{Total samples in Nyquist type II}}$)	4.76%	2.5%	1.72%	1.35%

rate 957.54 MHz) will be required since each CR requires 21 ADCs and there are 10 CRs in total. Thus, the system complexity of MS^3 is approximately half of that of the Nyquist system type I and much less than that of the Nyquist system type II.

In Fig. 6, we choose the CS-based system in [13] as a benchmark system due to its high impact and outstanding performance. In [13], the wideband spectrum was firstly modelled as a train of subbands which are smooth, but exhibit discontinuities or singularities at the subbands boundaries or edges. The wideband spectrum sensing was formulated as a standard CS problem. Basis pursuit-based optimisation algorithms were then applied to estimate the average frequency response amplitude of each subband, by which each subband can be coarsely detected. The comparison between the proposed MS^3 system and the benchmark system is provided. We assume that $v = 22$ CRs are collaborating for wideband spectrum sensing in both systems, in order to increase the reliability of spectrum sensing by exploiting spatial diversity. We see from Fig. 6(a) that MS^3 outperforms the CS-based system for every compression rate. In Fig. 6(b), it is seen that, compared with the benchmark system, MS^3 has better compression capability. Using MS^3, the probability of successful sensing becomes larger than 90% when the compression rate $\frac{M}{N} \geq 0.023$. In contrast, the benchmark system can achieve the probability of successful sensing at 90% only when the compression rate satisfies $\frac{M}{N} \geq 0.045$.

Furthermore, as shown in Table II, we find that the computational complexity of MS^3 is $\mathcal{O}(N \log N)$ due to energy detection with FFT operations, rather than $\mathcal{O}\left(N(M + \log N)\right)$ in the CS-based system, where M is usually much larger than $\log N$. The complexity of the CS-based system is hurt by both the matrix multiplication operations and the FFT operations for spectral recovery. To summarize, with the same computational resources, MS^3 has a relatively smaller spectrum sensing overhead than the CS-based system, due to not only the better compression capability (less data transmissions result in shorter transmission time), but also the lower computational complexity.

VI. CONCLUSIONS

In this paper, we have presented a novel system, MS^3, for cooperative wideband spectrum sensing in CR networks. MS^3 can relax the wideband spectrum sensing requirements of CRs due to its capability of sub-Nyquist sampling. It has been shown that, using sub-Nyquist samples, the wideband spectrum can be sensed in a cooperative manner without the need of spectral recovery, leading to a low spectrum sensing overhead. Moreover, we have derived closed-form bounds for the performance of MS^3 when combining faded or shadowed signals.

Simulation results have verified the derived bounds on the probabilities of false alarm and detection. It has also been shown that using partial measurements, MS^3 has superior performance even under low SNRs. The performance of MS^3 improves as the number of CRs or the average sampling rate increases. Compared with the existing wideband spectrum sensing methods, MS^3 not only provides computation and memory savings, but also reduces the hardware acquisition requirements and the energy costs at CRs.

APPENDIX A
RELATIONSHIP BETWEEN NYQUIST DFT SPECTRUM AND SUB-NYQUIST DFT SPECTRUM

Using the Poisson summation formula [22], (9), and (10), we obtain:

$$f_i \sum_{l\in\mathbb{Z}} X_{\mathrm{c},i}(w+f_i l) = \sum_{n\in\mathbb{Z}} y_i[n] e^{-\jmath 2\pi w n} = \sum_{n=0}^{M_i-1} y_i[n] e^{-\jmath 2\pi w n} = Y_i(w \tag{46}$$

where $X_{\mathrm{c},i}(w) = \int_{-\infty}^{\infty} x_{\mathrm{c},i}(t) e^{-\jmath 2\pi w t} dt$. Similar to (46), by using (1) and (2), we can obtain:

$$f \sum_{l\in\mathbb{Z}} X_{\mathrm{c},i}(w+fl) = \sum_{n\in\mathbb{Z}} x_i[n] e^{-\jmath 2\pi w n} = \sum_{n=0}^{N-1} x_i[n] e^{-\jmath 2\pi w n} = X_i(w) \tag{47}$$

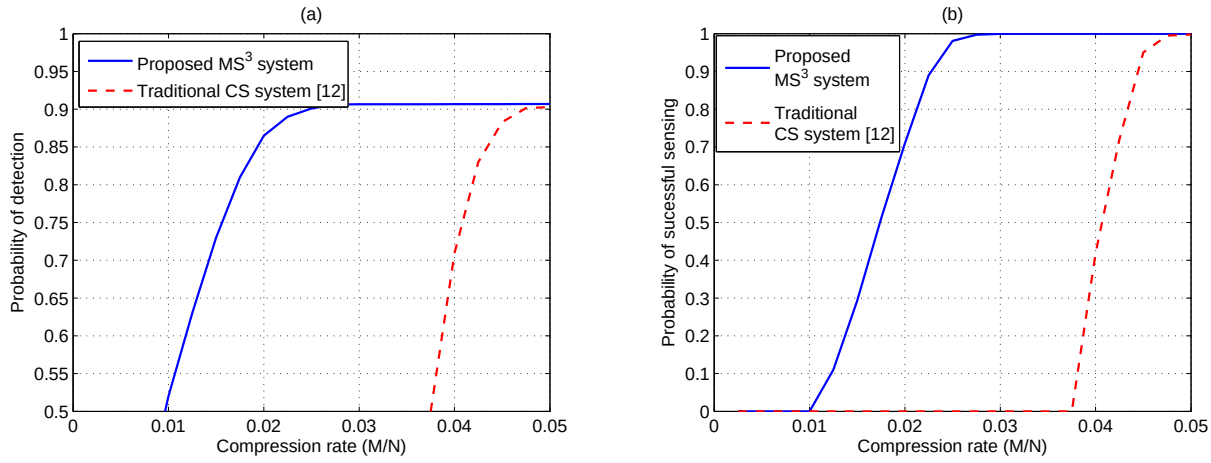

Fig. 6. Comparison between MS^3 and CS-based system [13]: (a) the probability of detection when the probability of false alarm is set to 10%, and (b) the probability of successful sensing which is defined as the probability of achieving both $P_{\rm d} \geq 90\%$ and $P_{\rm f} \leq 10\%$. In simulations, the average SNR as received at CRs is 10 dB and the number of CRs is $v = 22$.

TABLE II
COMPARISONS OF WIDEBAND SPECTRUM SENSING TECHNIQUES.

Approach	Compression Capability	ADC/DSP Type	Implementation Complexity	Computational Complexity
Wavelet detection	×	Nyquist	low	$\mathcal{O}(N \log N)$
Multiband joint detection	×	Nyquist	high	$\mathcal{O}(N \log N)$
CS-based detection	√	sub-Nyquist	medium	$\mathcal{O}(N(M + \log N))$
Proposed system	√	sub-Nyquist	low	$\mathcal{O}(N \log N)$

As the received signal is bandlimited and $f \geq 2W$, $X_i(w) = fX_{{\rm c},i}(w)$ holds for $w \in [-\frac{W}{2}, \frac{W}{2}]$. Substituting it to (46), we obtain $Y_i(w) = \frac{f_i}{f}\sum_{l=-\infty}^{\infty} X_i(w + f_i l)$. In a discrete form, we end up with:

$$Y_i[m] = \frac{M_i}{N}\sum_{l=-\infty}^{\infty} X_i[m + lM_i], \quad m = 0, 1, \cdots, M_i - 1. \tag{48}$$

APPENDIX B
PROBABILITY OF SIGNAL OVERLAP AT SUB-NYQUIST SAMPLING

As s spectral components are distributed over the frequency bins of $0, 1, \cdots, N-1$, the probability of the frequency bin k belonging to the spectral support Ω is $P = \Pr(k \in \Omega) = \frac{s}{N}$. Let q denote the number of spectral components overlapped on the frequency bin m, using (11) the probability of no signal overlap is given by

$$\Pr(q < 2) = \Pr(q = 0) + \Pr(q = 1) = (1-P)^{\lceil \frac{N}{M_i} \rceil} + \binom{\lceil \frac{N}{M_i} \rceil}{1} P(1-P)^{\lceil \frac{N}{M_i} \rceil - 1} \tag{49}$$

where $\lceil \frac{N}{M_i} \rceil$ denotes the number of summations in (11). Substituting $P = \frac{s}{N}$ into (49) while choosing sub-Nyquist sampling rate in MS^3 such that $M_i = \sqrt{N}$, we obtain

$$\Pr(q < 2) = \left(\frac{N-s}{N}\right)^{\frac{N}{M_i}} + \frac{s}{M_i}\left(\frac{N-s}{N}\right)^{\frac{N-M_i}{M_i}} = \frac{(\frac{N-s}{N})^{\sqrt{N}}(N - s + \sqrt{N})}{N - s} \tag{50}$$

It can be tested that $\Pr(q < 2)$ approaches to 1 when choosing N such that $N \gg s$. Thus the probability of signal overlap approaches to zero under the condition we choose.

APPENDIX C
PROOF OF LEMMA 1

Let M_i and M_j denote the number of samples at CRs i and j, respectively. Using (19), we can represent the aliased frequencies projected from $k_1 \in \Omega$ by

$$g_i = |k_1|_{\bmod (M_i)} + lM_i = k_1 - hM_i + lM_i, \ h \neq 0 \tag{51}$$
$$g_j = |k_1|_{\bmod (M_j)} + \check{l}M_j = k_1 - \check{h}M_j + \check{l}M_j, \ \check{h} \neq 0 \tag{52}$$

where integers h and $\check{h}$ are quotients from modulo operations, and $l - h \in [-\lceil \frac{N}{M_i} \rceil + 1, \lceil \frac{N}{M_i} \rceil - 1]$, $\check{l} - \check{h} \in [-\lceil \frac{N}{M_j} \rceil + 1, \lceil \frac{N}{M_j} \rceil - 1]$, in which $\lceil \frac{N}{M_i} \rceil$ gives the smallest integer not less than $\frac{N}{M_i}$.

Avoiding $g_i = g_j$ is equivalent to avoiding $(l - h)M_i = (\check{l} - \check{h})M_j$. If M_i and M_j are different primes, the condition $\max(|l - h|) < M_j$ (i.e., $\lceil \frac{N}{M_i} \rceil - 1 < M_j$) will satisfy this. After simplification, the condition $M_i M_j > N$ is obtained. Moreover, if it holds for any two CRs, the case for more than two CRs will also hold.

APPENDIX D
PROOF OF THEOREM 2

If the received signals at CRs are Rayleigh faded, the lower bound on the average probability of false alarm will remain as it is independent of the SNR. Using (25), (28), and (29),

the upper bound on the average probability of false alarm can be calculated by

$$\overline{P_{\mathrm{f},k}}^{\mathrm{up}} = \int_0^\infty Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{s}}}, \sqrt{\lambda_k}\right) \frac{\gamma_{\mathrm{s}}^{s-1}}{\overline{\gamma}^s \Gamma(s)} e^{-\frac{\gamma_{\mathrm{s}}}{\overline{\gamma}}} d\gamma_{\mathrm{s}}. \quad (53)$$

Rewriting the Marcum Q-function by using (4.74) in [28] and (8.352-2) in [29], we obtain:

$$Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{s}}}, \sqrt{\lambda_k}\right) = \sum_{n=0}^{\infty} \frac{\left(\frac{\psi\gamma_{\mathrm{s}}}{2}\right)^n e^{-\frac{\psi\gamma_{\mathrm{s}}}{2}}}{n!} \frac{\Gamma(n+Jv, \frac{\lambda_k}{2})}{\Gamma(n+Jv)}. \quad (54)$$

Substituting (54) into (53), we can rewrite (53) as

$$\overline{P_{\mathrm{f},k}}^{\mathrm{up}} = \frac{1}{\overline{\gamma}^s} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma(n+Jv, \frac{\lambda_k}{2})}{n!(s-1)!\Gamma(n+Jv)} \int_0^\infty \gamma_{\mathrm{s}}^{n+s-1} e^{-\frac{\psi\gamma_{\mathrm{s}}}{2} - \frac{\gamma_{\mathrm{s}}}{\overline{\gamma}}} d\gamma_{\mathrm{s}}. \quad (55)$$

Calculating the integral by using (3.351-3) in [29], we end up with

$$\overline{P_{\mathrm{f},k}}^{\mathrm{up}} = \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-s} \sum_{n=0}^{\infty} \mathrm{C}_{n+s-1}^{n} \left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^n \frac{\Gamma\left(n+Jv, \frac{\lambda_k}{2}\right)}{\Gamma\left(n+Jv\right)}. \quad (56)$$

Similarly, we can obtain the lower bound on the average probability of detection.

Appendix E
Proof of Theorem 3

If the received signals are shadowed according to log-normal distribution, the lower bound on $\widetilde{P_{\mathrm{f},k}}$ in (41) will remain. By (39), the upper bound on the probability of false alarm can be given by

$$\widetilde{P_{\mathrm{f},k}}^{u} = \int_0^\infty Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{s}}}, \sqrt{\lambda_k}\right) \sqrt{\frac{s\eta}{2\pi}} \gamma_{\mathrm{s}}^{-3/2} \exp\left(-\frac{\eta(\gamma_{\mathrm{s}} - s\theta)^2}{2s\theta^2\gamma_{\mathrm{s}}}\right) d\gamma_{\mathrm{s}} \quad (57)$$

Substituting (54) into (57), we calculate $\widetilde{P_{\mathrm{f},k}}^{u}$ as

$$\widetilde{P_{\mathrm{f},k}}^{u} = \sqrt{\frac{s\eta}{2\pi}} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma\left(n+Jv, \frac{\lambda_k}{2}\right)}{n!\Gamma\left(n+Jv\right)} \int_0^\infty \gamma_{\mathrm{s}}^{n-\frac{3}{2}} \exp\left(-\frac{s\psi\theta^2+\eta}{2s\theta^2}\gamma_{\mathrm{s}} + \frac{\eta}{\theta}\right) d\gamma_{\mathrm{s}}, \quad (58)$$

Using (3.471-9) in [29] for calculating the integral in (58), we obtain

$$\widetilde{P_{\mathrm{f},k}}^{u} = \sqrt{\frac{2s\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma\left(n+Jv, \frac{\lambda_k}{2}\right)}{n!\Gamma\left(n+Jv\right)} \left(\sqrt{\frac{s^2\eta\theta^2}{s\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}} \mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\eta\left(s\psi+\frac{\eta}{\theta^2}\right)}\right) \quad (59)$$

Likewise, the lower bound on the average probability of detection can be approximated.

References

[1] Z. Chen, C.-X. Wang, X. Hong, J. Thompson, S. Vorobyov, X. Ge, H. Xiao, and F. Zhao, "Aggregate interference modeling in cognitive radio networks with power and contention control," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 456–468, Feb. 2012.

[2] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.

[3] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tutorials*, vol. 11, no. 1, pp. 116 –130, Jan. 2009.

[4] D. Horgan and C. C. Murphy, "Fast and Accurate Approximations for the Analysis of Energy Detection in Nakagami-m Channels," *IEEE Commun. Letters*, vol. 17, no. 1, pp. 83–86, Jan. 2013.

[5] P.C. Sofotasios, E. Rebeiz, Li Zhang,T. A. Tsiftsis, D. Cabric, and S. Freear, "Energy Detection Based Spectrum Sensing Over $\kappa-\mu$ and $\kappa-\mu$ Extreme Fading Channels," *IEEE Trans. Vehicular Technology*, vol. 62, no. 3, pp. 1031–1040, 2013.

[6] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[7] S. P. Herath, N. Rajatheva, and C. Tellambura, "Energy Detection of Unknown Signals in Fading and Diversity Reception," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2443–2453, Sept. 2011.

[8] S. Atapattu, C. Tellambura, and Hai Jiang, "Performance of an Energy Detector over Channels with Both Multipath Fading and Shadowing," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3662–3670, Dec. 2010.

[9] S. Atapattu, C. Tellambura, and Hai Jiang, "Energy Detection Based Cooperative Spectrum Sensing in Cognitive Radio Networks," *IEEE Trans. Wireless Communications*, vol.10, no.4, pp.1232–1241, April 2011.

[10] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Sig. Processing*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[11] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Wideband spectrum sensing in cognitive radio networks," in *Proc. IEEE ICC'08*, Beijing, China, May 2008, pp. 901–906.

[12] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. IEEE CROWNC'06*, Mykonos Island, Greece, June 2006, pp. 1–5.

[13] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE ICASSP'07*, Hawaii, USA, April 2007, pp. 1357–1360.

[14] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-nyquist sampling for wideband spectrum sensing," *IEEE J. Sel. Topics Sig. Processing*, vol. 6, no. 1, pp. 58–69, Feb. 2012.

[15] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE J. Sel. Topics Sig. Processing*, vol. 5, no. 1, pp. 37–48, Feb. 2011.

[16] Y. Wang, Z. Tian, and C. Feng, "Sparsity order estimation and its application in compressive spectrum sensing for cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2116–2125, June 2012.

[17] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. IEEE ICASSP'09*, Taipei, April 2009, pp. 2337–2340.

[18] Y. Wang, A. Pandharipande, and G. Leus, "Compressive sampling based MVDR spectrum sensing," in *2010 2nd International Workshop on Cognitive Information Processing (CIP)*, June 2010, pp. 333–337.

[19] Y. Havary-Nassab, S. Hassan, and S. Valaee, "Compressive detection for wide-band spectrum sensing," in *in Proc. of IEEE ICASSP'10)*, Dallas, TX, USA, Mar. 2010, pp. 3094–3097.

[20] H. Sun, D. Laurenson, and C.-X. Wang, "Computationally tractable model of energy detection performance over slow fading channels," *IEEE Commun. Letters*, vol. 14, no. 10, pp. 924–926, Oct. 2010.

[21] H. Sun, W.-Y. Chiu, and A. Nallanathan, "Adaptive compressive spectrum sensing for wideband cognitive radios," *IEEE Commun. Letters*, vol. 16, no. 11, pp. 1812–1815, Nov. 2012.

[22] M. A. Pinsky, *Introduction to Fourier Analysis and Wavelets*. Providence, Rhode Island, USA: American Mathematical Society, 2002.

[23] P. Cantrell and A. Ojha, "Comparison of generalized Q-function algorithms," *IEEE Trans. Info. Theory*, vol. 33, no. 4, pp. 591–596, July 1987.

[24] C. W. Helstrom, "Approximate inversion of Marcum's Q-function," *IEEE Trans. Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 317–319, Jan. 1998.

[25] A. H. Marcus, "Power sum distributions: an easier approach using the Wald distribution," *J. of American Statistical Association*, vol. 71, pp. 237–238, 1976.

[26] R. S. Chhikara and J. L. Folks, *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*. New York: Marcel Dekker Inc., 1989.

[27] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics in Sig. Processing*, vol. 4, no. 2, pp. 375 –391, Apr. 2010.

[28] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. New York: John Wiley & Sons, Inc., Dec. 2004.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed., A. Jeffrey, Ed. New York: Academic Press, Inc., 1994.

Hongjian Sun [S'07-M'11] received his Ph.D. degree in 2010 at the University of Edinburgh, UK. He then joined King's College London, UK, as a Postdoctoral Research Associate in 2010. In 2011-2012, he was a visiting Postdoctoral Research Associate at Princeton University, USA. Since 2013, he has been a Lecturer in Smart Grid at the University of Durham, UK. His recent research interests include Smart Grids, Wireless Communications, and Signal Processing. He has made 1 contribution to the IEEE 1900.6a Standard, and published 2 book chapters and more than 50 papers in refereed journals and international conferences. He is on the Editorial Board for Journal of Communications and Networks, and EURASIP Journal on Wireless Communications and Networking, and was a Guest Editor for the special issue "Industrial Wireless Sensor Networks" for International Journal of Distributed Sensor Networks. Additionally, he is serving as an organizing chair for Workshop on Integrating Communications, Control, Computing Technologies for Smart Grid, Glasgow, UK, in May 2015, and Workshop on Communications Technologies for Smart Grid, Shanghai, China, in August 2015. He also served (or is serving) as a technical program committee (TPC) member for many international conferences, e.g., ICC, Globecom, VTC. He is a peer-reviewer for a number of international journals and was nominated as an Exemplary Reviewer by IEEE Communications Letters in both 2011 and 2012.

Arumugam Nallanathan [S'97-M'00- SM'05] is a Professor of Wireless Communications in the Department of Informatics at King's College London (University of London). He served as the Head of Graduate Studies in the School of Natural and Mathematical Sciences at King's College London, 2011/12. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include 5G Technologies, Millimeter wave communications, Cognitive Radio and Relay Networks. In these areas, he co-authored nearly 250 papers. He is a co-recipient of the Best Paper Award presented at the 2007 IEEE International Conference on Ultra-Wideband (ICUWB2007). He is a Distinguished Lecturer of IEEE Vehicular Technology Society.

He is an Editor for IEEE Transactions on Communications and IEEE Transactions on Vehicular Technology. He was an Editor for IEEE Transactions on Wireless Communications (2006-2011), IEEE Wireless Communications Letters and IEEE Signal Processing Letters. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society, Technical Program Co-Chair (MAC track) for IEEE WCNC 2014, Co-Chair for the IEEE GLOBECOM 2013 (Communications Theory Symposium), Co-Chair for the IEEE ICC 2012 (Signal Processing for Communications Symposium), Co-Chair for the IEEE GLOBECOM 2011 (Signal Processing for Communications Symposium), Technical Program Co-Chair for the IEEE International Conference on UWB 2011 (IEEE ICUWB 2011), Co-Chair for the IEEE ICC 2009 (Wireless Communications Symposium), Co-chair for the IEEE GLOBECOM 2008 (Signal Processing for Communications Symposium) and General Track Chair for IEEE VTC 2008. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.

Shuguang Cui [S'99-M'05-SM'12-F'14] received his Ph.D in Electrical Engineering from Stanford University, California, USA, in 2005. He has been working as an associate professor in Electrical and Computer Engineering at the Texas A&M University, College Station, TX. His current research interests focus on data oriented large-scale information analysis and system design, including large-scale distributed estimation and detection, information theoretical approaches for large data set analysis, complex cyber-physical system design, and cognitive network optimization. He was selected as the Thomson Reuters Highly Cited Researcher and listed in the Worlds Most Influential Scientific Minds by Sciencewatch in 2014. He was the recipient of the IEEE Signal Processing Society 2012 Best Paper Award. He has been serving as the TPC co-chairs for many IEEE conferences. He has also been serving as the associate editors for IEEE Transactions on Big Data, IEEE Transactions on Signal Processing, and IEEE Transactions on Wireless Communications. He was the elected member for IEEE Signal Processing Society SPCOM Technical Committee (2009 2014) and the elected Vice Chair for IEEE ComSoc Wireless Technical Committee (2015 2016). He is the member of the Steering Committee for the new IEEE Transactions on Big Data. He was elected as an IEEE Fellow in 2013 and an IEEE ComSoc Distinguished Lecturer in 2014.

Cheng-Xiang Wang [S'01-M'05-SM'08] received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, Jinan, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Aalborg, Denmark, in 2004.

Since 2005, he has been with Heriot-Watt University, Edinburgh, U.K., where he became a Professor in 2011. He is also an Honorary Fellow with the University of Edinburgh and a Chair/Guest Professor with Shandong University and Southeast University, Nanjing, China. He was a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005; a Visiting Researcher with Siemens AG-Mobile Phones, Munich, Germany, in 2004; and a Research Assistant with Technical University of Hamburg-Harburg, Hamburg, Germany, from 2000 to 2001. His current research interests include wireless channel modelling and 5G wireless communication networks, such as green communications, cognitive radio networks, high mobility communication networks, massive MIMO, millimeter wave communications, and visible light communications. He has edited one book, and published one book chapter and over 210 papers in refereed journals and conference proceedings.

Prof. Wang served or is currently serving as an Editor of eight international journals, including the IEEE Transactions on Vehicular Technology (2011-) and the IEEE Transactions on Wireless Communications (20072009). He was the leading Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Vehicular Communications and Networks. He served or is serving as a TPC member, a TPC Chair, and a General Chair for over 70 international conferences. He received the Best Paper Awards from IEEE Globecom 2010, IEEE ICCT 2011, ITST 2012, and IEEE VTC 2013-Fall. He is a Fellow of the IET, a Fellow of the HEA, and a member of EPSRC Peer Review College.